

Selfish Cops and Adversarial Robber: Multi-Player Pursuit Evasion on Graphs

G. Konstantinidis and Ath. Kehagias

March 23, 2017

Abstract

We introduce and study the game of “Selfish Cops and Adversarial Robber” (SCAR) which is an N -player generalization of the “classic” two-player cops and robbers (CR) game. We prove that SCAR has a Nash equilibrium in deterministic strategies.

1 Introduction

In this note we introduce and study the game of “*Selfish Cops and Adversarial Robber*” (SCAR) which is an N -player generalization of the “classic” two-player cops and robbers (CR) game [1, 8]. Our main result is that *SCAR has a Nash equilibrium in deterministic strategies*.

Here is a brief, informal description of SCAR (a rigorous definition will be given in the following sections). N players take turns moving their tokens along the edges of an undirected finite simple connected graph. $N - 1$ players are the cops and the remaining player is the robber. The robber is *captured* iff at the end of a turn he is located in the same vertex as a cop; in this case the capturing cop wins and the robber *as well as the remaining cops* lose. The robber wins if he is never captured.

What we have described above is the *qualitative*, win/lose variant of SCAR. But in this note we actually study a *quantitative* variant, in which each player is assigned a *payoff* dependent on the *capture time* T_C . Namely, we assume that the robber’s payoff is T_C , the capturing cop’s payoff is $-T_C$ and the non-capturing cop’s payoff is $-T_C - B$, where $B > 0$ is a “non-capture penalty”. So the robber tries to *maximize capture time*; each cop has a motive to minimize capture time and an additional motive for the capture to be effected by himself.

When $N = 2$ (i.e., one cop vs. one robber) SCAR reduces to the classical CR. When $N > 2$ we have a new game which, as far as we know, has not been previously studied. The novelty of the game lies in the fact that each cop plays *selfishly*; one cop’s win is another cop’s *partial* loss (as well as the robber’s complete loss).

In Section 2 we present the necessary preliminaries (notation etc.) for the analysis of three player (two cops, one robber) SCAR. In Section 3 we prove that three-player SCAR admits a Nash equilibrium in deterministic strategies. This result can be easily extended to N -player SCAR, as we explain in Section 4. In Section 5 we briefly discuss related work, possible extensions.

2 Preliminaries

We give a rigorous description of three-player SCAR (the extension to the N -player version is straightforward and will be discussed in Section 4). The game is played on a undirected finite simple connected graph $G = (V, E)$. The player set is $I = \{1, 2, 3\}$. The first (resp. second) player controls the first (resp. second) cop token; the third player controls the robber token (we will sometimes use the abbreviations C_1 , C_2 and R for both the players and their tokens).

The game is played in turns; in each turn a single player moves and the other players stay in place¹. A *game position* or *game state* is $s = (x^1, x^2, x^3, i)$ where $x^i \in V$ is the position (G vertex) of the i -th player and i is the number of the player who will make the next move. There is an additional *terminal* state which is denoted by \bar{s} . Hence the *set of all possible states* of the game is

$$S = (V \times V \times V \times I) \cup \{\bar{s}\}.$$

For each $i \in I$ we define the following sets of states

$$S^i = \{s = (x^1, x^2, x^3, i) : (x^1, x^2, x^3) \in V \times V \times V\},$$

i.e., S^i is the set of states in which the i -th player makes the next move. Also we define

$$\begin{aligned}\tilde{S}^1 &= \{s : s = (x^1, x^2, x^3, i) \text{ with } x^1 = x^3, x^2 \neq x^3\}, \\ \tilde{S}^2 &= \{s : s = (x^1, x^2, x^2, i) \text{ with } x^1 \neq x^3, x^2 = x^3\}, \\ \tilde{S}^{12} &= \{s : s = (x^1, x^2, x^2, i) \text{ with } x^1 = x^3, x^2 = x^3\}.\end{aligned}$$

So \tilde{S}^i is the set of C_i *capture states* (R is captured by C_i) and \tilde{S}^{12} is the set of C_{12} *capture states* (R is captured by both C_1 and C_2).

The *capture time* is

$$T_C = \min \{t : x_t^1 = x_t^3 \text{ or } x_t^2 = x_t^3\}$$

and $T_C = \infty$ if no capture takes place. We define:

$$\begin{aligned}\text{the set of capture states : } S_C &= \tilde{S}^1 \cup \tilde{S}^2 \cup \tilde{S}^{12} \\ \text{the set of non-capture states : } S_{NC} &= S / (S_C \cup \bar{s})\end{aligned}$$

Hence the state set can be partitioned into noncapture states, capture states and terminal states:

$$S = S_{NC} \cup S_C \cup \{\bar{s}\}.$$

This partition corresponds to the temporal evolution of the game. If $0 < T_C < \infty$, then:

1. at the 0-th turn the game starts at some preassigned state $s_0 \in S_{NC}$;
2. at the t -th turn (for $0 < t < T_C$), the game moves to some state $s_t \in S_{NC}$;
3. at the T_C -th turn the game moves to some capture state $s_T \in S_C$ and

¹The players' sequence is the obvious one: 1 is followed by 2 who is followed by 3 who is followed by 1 and so on.

4. at $t = T_C + 1$ the game moves to the terminal state and stays there: for every $t > T_C$, $s_t = \bar{s}$.

If, on the other hand, $T_C = \infty$ then $s_t \in S_{NC}$ for every $t \in \mathbb{N}_0$.

State to state transitions are effected (in the obvious manner) by the players' *moves* or *actions*. When the game state is $s = (x^1, x^2, x^3, j)$, the i -th player's *action set* is denoted by $A_i(s)$ and defined by

$$A_i(s) = \begin{cases} N[x^i] & \text{when } s \in S_i \cap S_{NC} \\ \{x^i\} & \text{when } s \in S_j \cap S_{NC} \text{ with } i \neq j \\ \emptyset & \text{when } s \in S_C \cup \{\bar{s}\} \end{cases}.$$

In other words, when the i -th player has the move, he can move to any vertex in the closed neighborhood of x^i ; when another player has the move, the i -th player can only stay in place (trivial move); when the game is in the terminal state or a capture state, the player has no moves.

According to the above rules, the game starts at some preassigned state $s_0 = (x_0^1, x_0^2, x_0^3, i_0)$ and at the t -th turn ($t \in \mathbb{N}$) is in the state $s_t = (x_t^1, x_t^2, x_t^3, i_t)$. This results in a *game history* $\mathbf{s} = s_0 s_1 s_2 \dots$. In other words, we assume each play of the game lasts an infinite number of turns; however, if $T_C < \infty$ then $s_t = \bar{s}$ for every $t > T_C$ (the game effectively ends at T_C). We define the following history sets:

$$\begin{aligned} \text{Histories of length } n : H_n &= \{\mathbf{s} = s_0 s_1 \dots s_n\}, \\ \text{Histories of finite length} : H_* &= \cup_{n=1}^{\infty} H_n, \\ \text{Histories of infinite length} : H_{\infty} &= \{\mathbf{s} = s_0 s_1 \dots s_n \dots\}. \end{aligned}$$

Using the standard game theory approach, we assume that at the start of the game the i -th player selects a *strategy* σ^i which determines all his subsequent moves. Because SCAR is a game of *perfect information*², a player loses nothing by using only *deterministic strategies*, i.e., functions

$$\sigma^i : H_* \rightarrow V$$

which assign a move to each finite-length history. We assume every player uses only *legal* strategies, i.e., ones which never produce moves outside the player's action set.

A *strategy profile* is a triple $\sigma = (\sigma^1, \sigma^2, \sigma^3)$. We define $\sigma^{-i} = (\sigma^j)_{j \in I \setminus \{i\}}$, e.g., $\sigma^{-1} = (\sigma^2, \sigma^3)$. A *stationary* strategy (also called a *positional* strategy) produces its next move depending only on the current state of the game (but neither previous states nor current time):

$$\sigma^i(s_0 s_1 \dots s_t) = \sigma^i(s_t).$$

To complete the definition of three-player quantitative SCAR, we define each player's *payoff function* so as to capture the idea presented in Section 1. For a given infinite history $\mathbf{s} = s_0 s_1 s_2 \dots \in H_{\infty}$ we can express $Q_i(\mathbf{s})$, the total payoff to the i -th player, as the sum of the turn payoffs

$$\forall i \in I : Q^i(\mathbf{s}) = \sum_{t=0}^{\infty} q^i(s_t)$$

²I.e., only one player moves at a time and each player knows every action of the players that moved before him at every point.

and the turn payoffs for each player are defined as follows.

1. The *robber* receives a positive payoff of 1 for every turn in which he is not captured:

$$q^3(s) = \begin{cases} 1 & \text{if } s \in S_{NC} \\ 0 & \text{otherwise.} \end{cases}$$

2. The i -th cop receives a negative payoff (penalty) of 1 for every turn in which the robber is not captured and a negative payoff of B for the turn (if any) in which the robber is captured by the other cop:

$$\forall i \in \{1, 2\} : q^i(s) = \begin{cases} -1 & \text{if } s \in S_{NC} \\ -B & \text{if } s \in S_{C_j} \text{ with } j \neq i \\ 0 & \text{else.} \end{cases}$$

Given a strategy profile $\sigma = (\sigma^1, \sigma^2, \sigma^3)$ and an initial state s , the game history is fully determined. So we can denote the payoffs by $Q_s^i(\sigma)$, $Q_s^i(\sigma^1, \sigma^2, \sigma^3)$ or $Q_s^i(\sigma^j, \sigma^{-j})$.

3 Three Player SCAR

We will now prove that three-player SCAR has a Nash equilibrium in deterministic strategies. To this end we will use an approach which has previously been used for several other N -player games of perfect information [3, 4, 9], namely the use of *threat strategies*.

To this end we first introduce the auxiliary games $\Gamma_s(i)$ with $s \in S_{NC}$ and $i \in I$. The game $\Gamma_s(i)$ is the *two-player zero-sum* game with initial state s and played by player i (with payoff Q^i) against the *coalition* of players $I \setminus \{i\}$ (with payoff $-Q^i$).

So, for example, $\Gamma_s(3)$ is the classical CR game played by a single robber against two cooperating cops and with initial state s ; the robber's payoff (and the cops' penalty) is the capture time T_C . Similarly, $\Gamma_s(1)$ is a CR game (with initial state s) in which a single cop plays against the team of a robber and a cooperating cop; the single cop's penalty is the capture time T_C which is also the payoff to the robber-cop team³.

Using the terminology and results of [5], for every $s \in S_{NC}$ we have that $\Gamma_s(3)$ is a two-player, zero-sum *positive* stochastic game and hence has a (minimax) *value*, the *minimizer* (cop team) has an optimal stationary strategy and the *maximizer* (robber) has an ε -optimal stationary strategy. Since $\Gamma_s(3)$ is a perfect information game with integer payoffs the maximizer actually has a 0-optimal (i.e., optimal) strategy as well and all strategies are deterministic and stationary. Similarly, $\Gamma_s(1)$ (resp. $\Gamma_s(2)$) is a positive stochastic game in which the minimizer is C_1 (resp. C_2) and the maximizer is the coalition of C_2 and R (resp. C_1 and R). Hence standard results [5, Theorem 4.4.3] give the following.

Lemma 1 *For each $s \in S_{NC}$ and $i \in I$, the game $\Gamma_s(i)$ has a value and the players have deterministic and stationary optimal strategies.*

³To further clarify, in a graph with cop number 1, an optimal strategy for the robber-cop team is that they meet in the longest possible time but before the robber is caught by the first cop. In a graph with cop number grater than 1, an optimal strategy is for the second cop to always avoid the robber and the robber to always avoid both cops.

For each game and each player, the value and optimal strategies can be computed by the value-iteration algorithm provided in [5, Theorem 4.4.4]; this algorithm reduces to the algorithm of [6] for $\Gamma_s(3)$ and very similar algorithms for $\Gamma_s(1)$ and $\Gamma_s(2)$. Adapting the notation of [9] let us denote by ϕ_i^i the optimal (maxmin) strategy of player i in $\Gamma_s(i)$ and by ϕ_i^{-i} the joint (optimal) strategy of the coalition I/i against i (so the player $j \in I/i$ has the optimal strategy ϕ_i^j)⁴.

Returning to the three-player game, the previously mentioned threat strategy of player i is denoted by π^i and “composed” from the ϕ_i^j ’s as follows:

1. as long as every player $j \neq i$ follows ϕ_j^j , player i follows ϕ_i^i ;
2. as soon as some player $j \neq i$ deviates from ϕ_j^j , player i switches to ϕ_i^{-i} and uses it for the rest of the game.

In other words, the threat strategies in the three-player game are the following. Each player i plays the strategy which is optimal in $\Gamma_s(i)$, as long as the other players do the same. If at some point player j deviates from the above⁵, then the players in $I \setminus \{i\}$ form a coalition and play the coalition strategy against which is optimal against j in $\Gamma_s(j)$.

Our main result is the following.

Theorem 2 *In three-player SCAR, for any starting state s we have*

$$\forall i \in I, \forall \sigma^i : Q_s^i(\pi^1, \pi^2, \pi^3) \geq Q_s^i(\sigma^i, \pi^{-i}). \quad (1)$$

In other words, $\pi = (\pi^1, \pi^2, \pi^3)$ is a Nash equilibrium for three-player SCAR with any starting state s .

Proof. We choose some initial state s and fix it for the rest of the proof. Now let us prove (1) for the case $i = 1$. In other words, we need to show that

$$\forall \sigma^1 : Q_s^1(\pi^1, \pi^2, \pi^3) \geq Q_s^1(\sigma^1, \pi^2, \pi^3). \quad (2)$$

We take any σ^1 and define

$$\begin{aligned} \widehat{\mathbf{s}} &= \widehat{s}_0 \widehat{s}_1 \widehat{s}_2 \dots \text{ is the history produced by } (\pi^1, \pi^2, \pi^3), \\ \widetilde{\mathbf{s}} &= \widetilde{s}_0 \widetilde{s}_1 \widetilde{s}_2 \dots \text{ is the history produced by } (\sigma^1, \pi^2, \pi^3). \end{aligned}$$

Define

$$T_1 = \min \{t : \widetilde{s}_t \neq \widehat{s}_t\},$$

i.e., T_1 is the earliest time in which σ^1 produces a different state from π^1 . If $\widetilde{\mathbf{s}} = \widehat{\mathbf{s}}$, then $T_1 = \infty$ and

$$Q_s^1(\pi^1, \pi^2, \pi^3) = Q^1(\widehat{\mathbf{s}}) = Q^1(\widetilde{\mathbf{s}}) = Q_s^1(\sigma^1, \pi^2, \pi^3). \quad (3)$$

⁴Note that, since these strategies are stationary they do not depend on the starting state s ; in fact the same ϕ_i^j is optimal for every game $\Gamma_s(i)$.

⁵Note that this deviation will be detected immediately, since the game has perfect information.

If $T_1 < \infty$, then at $t = T_1 - 1$ player 1 deviated from ϕ_1^1 ; at $t = T_1$ the deviation was detected and player 2 (resp. 3) who still uses π^2 (resp. π^3) switched to ϕ_1^2 (resp. ϕ_1^3). It follows that

$$Q_s^1(\pi^1, \pi^2, \pi^3) = Q^1(\hat{\mathbf{s}}) = \sum_{t=0}^{T_1-2} q^1(\hat{s}_t) + \sum_{t=T_1-1}^{\infty} q^1(\hat{s}_t), \quad (4)$$

$$Q_s^1(\sigma^1, \pi^2, \pi^3) = Q^1(\tilde{\mathbf{s}}) = \sum_{t=0}^{T_1-2} q^1(\tilde{s}_t) + \sum_{t=T_1-1}^{\infty} q^1(\tilde{s}_t). \quad (5)$$

Since $\tilde{s}_t = \hat{s}_t$ for every $t < T_1$, it suffices to compare the second sums of (4) and (5). Let $s^* = \hat{s}_{T_1-1} = \tilde{s}_{T_1-1}$. Since $\phi_1^1, \phi_1^2, \phi_1^3$ are stationary, the history $\hat{s}_0 \hat{s}_1 \dots \hat{s}_{T_1-2}$ does not influence the moves they will produce at times $T_1, T_1 + 1, \dots$. Hence we have

$$\sum_{t=T_1-1}^{\infty} q^1(\hat{s}_t) = Q^1(\hat{s}_{T_1-1} \hat{s}_{T_1} \hat{s}_{T_1+1} \dots) = Q_{s^*}^1(\pi^1, \pi^2, \pi^3) = Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3) \geq Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3). \quad (6)$$

In other words, the sum in (6) is the payoff (to player 1) of the three-player game started from s^* with player i using strategy ϕ_i^i . But $Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3)$ is also the payoff of the two-player, zero-sum game $\Gamma_{s^*}(1)$, i.e., starting at s^* , with player 1 playing ϕ_1^1 and the coalition of players 2 and 3 playing (ϕ_1^2, ϕ_1^3) . However, we know that in $\Gamma_{s^*}(1)$ the optimal coalition strategy is (ϕ_1^2, ϕ_1^3) ; hence we have the inequality at the end of (6).

Next consider $\tilde{\mathbf{s}} = \tilde{s}_0 \tilde{s}_1 \tilde{s}_2 \dots$. It is produced by (σ^1, π^2, π^3) and, since σ^1 is not necessarily stationary, $\tilde{s}_{T_1} \tilde{s}_{T_1+1} \tilde{s}_{T_1+2} \dots$ could depend on $\tilde{s}_0 \tilde{s}_1 \dots \tilde{s}_{T_1-2}$. However, we can introduce the strategy ρ^1 induced by σ^1 on the subgame starting at s^* , which will produce the same history $\tilde{s}_{T_1} \tilde{s}_{T_1+1} \tilde{s}_{T_1+2} \dots$ as σ^1 .⁶ Then we have

$$\begin{aligned} \sum_{t=T_1-1}^{\infty} q^1(\tilde{s}_t) &= Q^1(\tilde{s}_{T_1-1} \tilde{s}_{T_1} \tilde{s}_{T_1+1} \dots) \\ &= Q_{s^*}^1(\rho^1, \pi^2, \pi^3) = Q_{s^*}^1(\rho^1, \phi_1^2, \phi_1^3) \leq Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3). \end{aligned} \quad (7)$$

The inequality in (7) is a consequence of the optimality of ϕ_1^1 as a response to ϕ_1^2, ϕ_1^3 in $\Gamma_s(1)$.

Combining (4)-(7) we have:

$$\begin{aligned} Q_s^1(\sigma^1, \pi^2, \pi^3) &= \sum_{t=0}^{T_1-2} q^1(\tilde{s}_t) + Q_{s^*}^1(\rho^1, \phi_1^2, \phi_1^3) \\ &\leq \sum_{t=0}^{T_1-2} q^1(\tilde{s}_t) + Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3) \\ &\leq \sum_{t=0}^{T_1-2} q^1(\hat{s}_t) + Q_{s^*}^1(\phi_1^1, \phi_1^2, \phi_1^3) \\ &= Q_s^1(\pi^1, \pi^2, \pi^3) \end{aligned}$$

⁶We define ρ^1 such that, when combined with $\tilde{s}_{T_1-1}, \phi_1^2, \phi_1^3$, will produce the same history $\tilde{s}_{T_1} \tilde{s}_{T_1+1} \tilde{s}_{T_1+2} \dots$ as σ^1 . Note that ρ^1 will in general depend (in an indirect way) on $\tilde{s}_0 \tilde{s}_1 \dots \tilde{s}_{T_1-2}$.

and we have proved (2), which is (1) for $i = 1$. The proof for the cases $i = 2$ and $i = 3$ are similar and hence omitted. ■

Before concluding this section it is worth emphasizing the following points.

1. A strategy profile $\pi = (\pi^1, \pi^2, \pi^3)$ forms a Nash equilibrium iff (for every i) player i has no incentive to *unilaterally* deviate from strategy π^i (i.e., while the players $I \setminus \{i\}$ stay with strategies π^{-i}). This does not imply any concept of global optimality. In other words, player i may be able to achieve a payoff higher than $Q^i(\pi)$ if more than one players deviate from the strategy profile π . A related point is that a game may possess more than one Nash equilibria. Hence 3-player SCAR may possess Nash equilibria different from the one indicated in Theorem 2 and these may in fact provide better payoff for one or more players.
2. It must be noted that the strategies (π^1, π^2, π^3) of Theorem 2 are *not* stationary. In particular, the action of a player at time t may be influenced by the action (deviation) performed by another player at time $t - 2$.
3. It seems reasonable (but we currently have no rigorous proof) that the solution of *qualitative* (win/lose) SCAR can be obtained by letting the non-capture penalty B tend to ∞ .

4 N -Player SCAR

The extension to N -player SCAR is straightforward. The auxiliary two-player zero-sum games are $\Gamma_s(1), \dots, \Gamma_s(N)$, where $\Gamma_s(i)$ is the two-player game with initial state s in which player i (with payoff Q^i) plays against the coalition $I \setminus \{i\}$ (with payoff $-Q^i$). Similarly to the 3-player case, for each $s \in S_{NC}$ and $i \in I$, the game $\Gamma_s(i)$ has a value and the players have deterministic and stationary optimal strategies. The meanings of ϕ_i^i and ϕ_i^{-i} are as in Section 3. The threat strategy of player i in the N -player game is π^i defined (exactly as in Section 3) as follows:

1. as long as every player $j \neq i$ follows ϕ_j^j , player i follows ϕ_i^i ;
2. as soon as some player $j \neq i$ deviates from ϕ_j^j , player i switches to ϕ_i^{-i} and uses it for the rest of the game.

Using the same approach as in Section 4 we can obtain the following.

Theorem 3 *In N -player SCAR, for any starting state s we have*

$$\forall i \in I, \forall \sigma^i : Q_s^i(\pi) \geq Q_s^i(\sigma^i, \pi^{-i}). \quad (8)$$

In other words, $\pi = (\pi^1, \pi^2, \dots, \pi^N)$ is a Nash equilibrium for N -player SCAR with any starting state s .

5 Concluding Remarks

In this note we have introduced SCAR, an N -player cops and robber game. As far as we know, games of this type have not been previously studied (but see [7] for a game between *two* selfish cops). We emphasize that SCAR involves several *independently and selfishly acting* cops, so it is different from previously introduced games involving several cops controlled by a single player [6].

In proving that SCAR has a Nash equilibrium we have used *threat strategies*, i.e. strategies by which any player's deviation is immediately punished by a coalition of the remaining players (essentially reverting to a two-player game). As mentioned, threat strategies have been used to solve several other classes of N -player games [3, 4, 9] but, as far as we have known, have not been previously applied to cops and robber games. In fact it is straightforward to generalize the use of threat strategies to N -player *generalized pursuit games* (for some generalized *two*-player pursuit games see [2]).

References

- [1] Bonato, Anthony, and Richard J. Nowakowski. "The game of cops and robbers on graphs". Vol. 61. Providence: American Mathematical Society, 2011.
- [2] Bonato, Anthony, and Gary MacGillivray. "Characterizations and algorithms for generalized Cops and Robbers games", accepted to *Contributions to Discrete Mathematics*.
- [3] Boros, Endre, and Vladimir Gurvich. "Why chess and back gammon can be solved in pure positional uniformly optimal strategies". Rutcor Research Report 21-2009, Rutgers University, 2009.
- [4] Chatterjee, Krishnendu, Rupak Majumdar, and Marcin Jurdziński. "On Nash equilibria in stochastic games". In International Workshop on Computer Science Logic, pp. 26-40. Springer Berlin Heidelberg, 2004.
- [5] Filar, Jerzy, and Koos Vrieze. *Competitive Markov decision processes*. 1996.
- [6] Hahn, Geňa, and Gary MacGillivray. "A note on k-cop, l-robber games on graphs." *Discrete mathematics* 306, no. 19 (2006): 2492-2497.
- [7] Kehagias, Athanasios, and Georgios Konstantinidis. "Selfish Cops and Passive Robber: Qualitative Games". arXiv preprint arXiv:1607.05434 (2016).
- [8] Nowakowski, Richard, and Peter Winkler. "Vertex-to-vertex pursuit in a graph". *Discrete Mathematics* 43, no. 2-3 (1983): 235-239.
- [9] Thuijsman, Frank, and Thirukkannamangai ES Raghavan. "Perfect information stochastic games and related classes". *International Journal of Game Theory* 26, no. 3 (1997): 403-408.